WIX3001 Soft Computing

Assignment 1: MATLAB Programming

For this assignment, refer to Tutorial 4. The tutorial shows how to use genetic algorithms (GA) to optimize the weights of a multi-layer perceptron (MLP). Use the tutorial as a starting point for this assignment.

1. Find at least three other datasets. Examples can be found from https://archive.ics.uci.edu/ml/datasets.php. Prepare the datasets similar to "iris.csv".

2. The goal is to optimize a MLP for each dataset.
    a. **Exercise 1**: modify "***create_network.m***" so that the parameters "***input_layer_units***", "***hidden_layers***", "***hidden_layer_units***", and "***output_layer_units***" can be changed from "***run_ann_withGA.m***". The number of input and output units should match whichever dataset you are using.

3. Run GA to optimize the MLP weights for the new dataset. You can set the population size and maximum generations to any number you want.
    a. **Exercise 2**: record the progress of the GA, specifically the average fitness and maximum fitness. Plot both fitness scores as a graph vs number of generations, and include the plots in the final report.
    b. **Exercise 3**: in addition to optimizing MLP weights, modify the GA to also optimize "***hidden_layers***" and "***hidden_layer_units***".
    c. **Exercise 4**: for each dataset, run at least 3 times, each time using a different RNG seed number.

4. After the GA finishes running, save the results into a separate .mat file for each dataset and for each seed number. For example: the results (the variables ***population***, ***fitness_score***, and ***progress***) for the Iris dataset for the seed number 1 should be saved as "iris_1.mat".

5. Report your results in a Word document.
    a. Describe which datasets you used: how many samples in the dataset, how many features, and how many classes.
    b. Report the results in a table. For example:

		Generation = 1		Generation = Last Generation			
	SeedNumber	AvgFitness	MaxFitness	AvgFitness	MaxFitness	Number of Layers *Report average and standard deviation based on the final population.	Number of Units for each Layer *Report average and standard deviation based on the final population
Dataset_1	1						
	2						
	3						
Dataset_2	4						
	5						
	6						
Dataset_3	7						
	8						
	9						

6. Zip your report, the MATLAB codes, and the results files and upload to UM Spectrum by **12th May 2023**.